



Title	Interaction between lexical tone and labial movement in Cantonese bilabial plosive production
Author(s)	Tong, Tik-sang
Citation	
Issued Date	2011
URL	http://hdl.handle.net/10722/192906
Rights	The author retains all proprietary rights, (such as patent rights) and the right to use in future works.

Interaction between lexical tone and labial movement in Cantonese bilabial
plosive production

UID: 2007250719

A dissertation submitted in partial fulfilment of the requirements for the Bachelor of Science
(Speech and Hearing Sciences), The University of Hong Kong, June 30, 2011.

Interaction between lexical tone and labial movement in Cantonese bilabial
plosive production

Eric Tong Tik Sang

A dissertation submitted in partial fulfilment of the requirements for the Bachelor of Science
(Speech and Hearing Sciences), The University of Hong Kong, June 30, 2011.

Abstract

The present study investigated the possible laryngeal-articulatory interaction by examining the direct kinematic data of Cantonese bilabial plosive produced at different lexical tones. Aspirated and unaspirated bilabial plosives /p^h/ and /p/ produced at six lexical tones (T1 - T6) by three native adult male speakers of Cantonese were studied to reveal the effect of tone on articulatory gestures.

By using electromagnetic articulography (EMA), movement of the jaw, upper lip and lower lip was measured. The kinematic data revealed that high level tone (T1) was associated with a considerably faster lip movement than low level tone (T6) in terms of lip opening velocity (LOV) and time to peak velocity of lip opening (TPV). This finding was in line with results previously reported on Mandarin (Hoole & Hu, 2004). The high subglottal pressure required by high level tone production can be used to explain the faster lip movement during bilabial separation.

However, unaspirated plosives of a lower intraoral pressure exhibited faster lip movements (as indicated by LOV, lip peak velocity and TPV values) during lip opening than aspirated plosives. Further investigations in aerodynamics and laryngeal activities are suggested in order to draw a more concrete conclusion for aspiration effect.

Introduction

The source-filter theory has been used to explain the mechanism of human speech sound production: the speech output is a product of the acoustic signal generated by vocal fold vibrations (glottal signal) being modulated by the resonance characteristics of the supralaryngeal vocal tract (transfer function) (Pickett, 1998). In the derivation of this theory, independence of the laryngeal (source) and supralaryngeal (filter) systems is assumed. According to Fant (1960), the vocal tract was depicted as four homogenous tubes with one end open and another end closed. These four tubes are the buccal cavity, front cavity, tongue section and back cavity. With variations in length and cross section area of the tubes, different speech sounds are produced with resonances at specific formant frequencies; yet output from the larynx is unchanged (Fant, 1960). There is no interaction between the source and the filter during speech production. However, in recent acoustic and kinematic studies, the validity of this assumption has been challenged.

In a study of vowel production by alaryngeal speakers, Ng and Chu (2009) found that articulation of vowels was changed after removal of the larynx in laryngectomized individuals. This suggested that alternation of the “source” could change the articulatory behavior of the “filter” during speech production. The independence relationship between laryngeal and supralaryngeal systems may be invalid. Interaction between the laryngeal system and supralaryngeal system appears to exist.

Tracking of the articulatory gestures can provide essential kinematic data that can help examine the possible interaction of supralaryngeal and laryngeal systems. Traditionally, movement information of speech articulators could be obtained by using cineradiography or x-ray techniques. But this was proven to have potential hazards and might complicate the direct

viewing and tracking of articulatory movements during continuous speech (Perkell et al., 1992). In recent years, a new technology - electromagnetic articulography (EMA) has gained its popularity. EMA is a biologically safe, non-invasive technique that can be used to examine the real-time, three-dimensional articulatory movements (Zierdt, Hoole, & Tillmann, 1999) and obtain first-hand kinematic data without interfering speech movements (Goozée, Murdoch, Theodoros, & Stokes, 2000). With these advantages, it has been used to assess the articulatory movement for both medical and research purposes. For example, EMA has been used to assess articulatory coordination in dysarthric patients after traumatic brain injuries (Bartle, Goozée, Scott, Murdoch, & Kuruvilla, 2006) and in comparing lingual kinematics for differentiated and undifferentiated lingual movements of speech-disordered children (Goozée et al., 2007). Furthermore, EMA technology has also promoted studies testifying the independence relationship between the laryngeal and supralaryngeal systems (Hoole, Zierdt, & Geng, 2003).

With the EMA technique, studies on the supralaryngeal and laryngeal interaction by tonal production have been done in various languages, including Mandarin. For example, Hoole and Hu (2004) studied articulatory gestures of vowels /i/, /y/, /u/, /e/ and /a/ produced at different Mandarin lexical tones. In the study, EMA was used to track the movement of jaw, lips and tongue during speech production. They found a significant effect of tone production on vowel articulation in terms of the position of jaw, lips and tongue (Hoole & Hu, 2004). Different lexical tones were associated with slightly different articulatory gestures. This finding provides insight about the relationship between the laryngeal system (represented by different lexical tones) and the supralaryngeal system (represented by varying articulatory gestures). In another study investigating the relationship between lexical tone and articulation of Mandarin syllables, Erickson, Iwata, Endo and Fujino (2004) examined the movements of tongue and jaw when the

syllables /ba/, /ma/ and /pa/ were produced at two lexical tones (Tone 1, high tone and Tone 3, low falling-rising tone). A significant difference was found in tongue and jaw movement between the two tones during speech production. The results indicated the presence of interactive patterns the articulatory gesture and tone variations. In other words, there was a correlation between our laryngeal and supralaryngeal systems.

Though findings from these studies quite consistently contradicted the assumption of the source-filter theory, further examination based on other tone languages is still needed to confirm this as previous studies were based only on Mandarin. It is not known if this is also true for other languages. Meanwhile, previous studies only measured jaw and tongue movements (Erickson et al., 2004). Kinematic information from other articulators such as soft palate and the lips is not available. To further the kinematic investigation in tone languages, the present study attempted to examine the possible interaction between laryngeal and supralaryngeal systems by examining the possible relationship between labial movements in bilabial plosive produced at different Cantonese lexical tones.

In Cantonese, aspirated and unaspirated bilabial plosives /p^h/ and /p/ are contrastive, meaning that, for the same phonetic context, if an aspirated bilabial plosive is substituted by an unaspirated counterpart, there will be a change in meaning. Although both aspirated and unaspirated bilabial plosives have the same place of articulation (bilabial), subtle difference may exist between them. Study of possible aspiration effect on the specific articulation of bilabial plosives will shed light on the understanding of fine motor control during production of different plosives. In a kinematic study of English whispered bilabial plosives, Higashikawa, Green, Moore, and Minifie (2003) found the significant difference in lip movement between whispered /b/ and whispered /p/, providing an insight about the possible aspiration effect on Cantonese

bilabial plosive production. Therefore, investigating the aspiration effect on articulation was also an aim of the present study.

During bilabial plosive production, a complete labial closure is formed and released rapidly. According to Clark, Yallop, and Fletcher (2007), the entire process of a bilabial plosive production can be divided into three phases: (1) occlusion phase, (2) hold phase, and (3) release phase. Firstly, formation of a lip closure is completed during the occlusion phase. It is followed by the building up of a positive intraoral pressure which is maintained behind the oral occlusion during the hold phase. As production of Cantonese bilabial plosives is egressive, an outward flow of airstream is produced. Finally, the labial occlusion is released rapidly by the forced opening of lips due to the high intra-oral pressure, i.e., release burst in release phase (Clark, Yallop, & Fletcher, 2007). It should be noted that the release phase is followed by the onset of vocal fold vibration, which corresponds to phonation of the following vowel. Movement of the two lips is usually visible to the listener. This, in fact, serves as an important visual cue for the correct identification of bilabial plosives.

According to the source-filter theory, lip movement (supralaryngeal system) during production of bilabial plosive should be independent of vocal fold phonation (laryngeal system). Vocal fold behavior should not affect, or be affected by, the release of stop in the supralaryngeal vocal tract. In this study, lip movement of the same syllable formed by an initial bilabial plosive and the vowel /a/ but produced at different Cantonese lexical tones was examined. The main research question was: *Do lip movements change with variations in tones in Cantonese, in a way similar to what was observed in Mandarin?* In addition, the present study also examined the possible aspiration effect on articulation of bilabial plosives, and if there is any interaction

between aspiration and tone. Findings from the present study helped answer the question about the laryngeal-supralaryngeal independence relationship assumed by the source-filter theory.

Methods

Participants

Three adult male native Cantonese speakers (aged from 19 to 23 years) participated in the study. A relatively small subject pool was used as EMA study is time-consuming and laborious. Yet, such sample size was comparable to many published studies in the literature (e.g., Goozée, Murdoch, Theodoros, & Stokes, 2000; Hoole & Hu, 2004). All participants were healthy speakers with no history of neurological disorders, oro-maxillo-facial surgery, and/or speech, language, or hearing problems. Contraindications of the instrument were avoided such as use of pace maker, claustrophobia and electromagnetic hypersensitivity. Only those who could speak naturally with the sensor coils attached were recruited.

Speech Material

The speech materials consisted of the monosyllabic Cantonese bilabial words /pa/ and /p^ha/ produced at the six Cantonese lexical tones: Tone 1 to Tone 6 (T1-T6). The six Cantonese tones represented by T1 to T6 are the high-level, high-rising, mid-level, low-falling, low-rising and low-level tones. During the experiment, Chinese characters corresponding to the monosyllables were printed on a sheet that was provided to the participants. For the non-word /pa⁵/, /p^ha⁵/ and /p^ha⁶/, demonstrations were given to the participants and practice was given until they could be produced correctly by the participants.

EMA Preparation

A three-dimensional EMA system (AG500, CarstensMedizinelektronik) was used to measure the movement of upper and lower lips, and lower jaw during production of the

Cantonese monosyllables. Warming up and calibration of the EMA system, and preparation and attachment of sensors were done strictly following the procedures described in the AG500 system manual before data collection.

The EMA system was comprised of three pairs of transmitter coils which were installed on a transparent plastic cube. During the experiment, the transmitter coils generated an alternating magnetic field of frequency from 7.5 kHz to 13.75 kHz and induced alternating signals in the small receiver coils (2 x 2 x 3mm) which were attached on the articulators of participants upon movement. Receiver coils coated with latex were affixed to the lower jaw, upper and lower lips using biologically safe adhesive. Additionally, three receiver coils were attached to fixed reference points for head movements, i.e. the bridge of nose and the mastoid processes of the left and right temporal bones.

The magnitude of the induced signals in receiver coils was derived from the distance between the transmitter and transducer. Hence, movement of lips and jaw in y-z coordinate plane was obtained by measuring the position of receiver coils within a period of time. The sampling frequency was set at 200Hz (200 samples per second). Displacement (mm), velocity (mm/s) and duration (s) data were then calculated using the magnitude of the induced signals.

Recordings

At the time of experiment, the participant was positioned in the center of the EMA cube with the sensor coils securely attached to the designated positions. They were instructed to produce 48 Cantonese monosyllabic words (2 syllables x 6 tones x 4 repetitions) at a comfortable loudness level and speech rate. To eliminate possible order effect, the order at which the stimuli were produced was randomized. Before the actual recording took place, the participants were given a five-minute practice period to accustom themselves to the recording environment and

speaking with receiver coils attached. Recording began only when participants were able to produce the speech sounds naturally with no complaint from the participants. For referencing, acoustic data were recorded simultaneously. The entire recording lasted for about 30 minutes for each participant.

Data Analyses

Movement of the lower jaw, and upper and lower lips were measured during production of aspirated and unaspirated plosives /p^h/ and /p/ at different Cantonese tones. However, since changes in signal magnitude of receiver coils could also be contributed by head movement, correction for possible head movements relative to the EMA cube was needed. This was done to reposition movement data by taking references from the receiver coils on the bridge of nose and the two temporal bones. Kinematic data of the articulators included distance (mm), velocity (mm/s), and duration (s) of the two lips. Only kinematic data in the z-axis was analyzed as it contributes to the change in vertical dimension of oral opening and hence the pitch variation most. Mean values of each kinematic parameter were calculated by averaging four productions.

Movement data were grouped by the two independent variables: aspiration (aspirated vs. unaspirated) and tones (T1-T6). Lip opening velocity (mm/s), peak velocity of lip movement (mm/s), and time-to-peak velocity of lip movement(s) during production of /p^h/ and /p/ were compared.

Lip opening velocity (LOV) is defined as the average rate at which the lips open during the burst phase of bilabial plosive production. It corresponds to the time required to achieve the greatest lip opening. The shorter is the time for reaching the greatest lip opening, the higher is the LOV value. Figure 1 depicts the lip separation with time of 12 plosive productions in which each wave represents one production. Lip separation is calculated as the difference between the

maximum and minimum height. In the study, LOV of each plosive was obtained by dividing the maximum separation of the two lips by the time elapsed.

Peak velocity is used to indicate the greatest velocity of lip movement during the burst phase of bilabial plosive production. As z-coordinate was concerned, the peak velocity measured was recognized for the vertical lip opening movement. Accordingly, time-to-peak velocity (TPV) is defined as the time required for reaching the highest velocity of lip opening movement during the burst phase of bilabial plosive production. In the present study, kinematic data of bilabial movements were compared among different tones and between aspirated and unaspirated plosives.

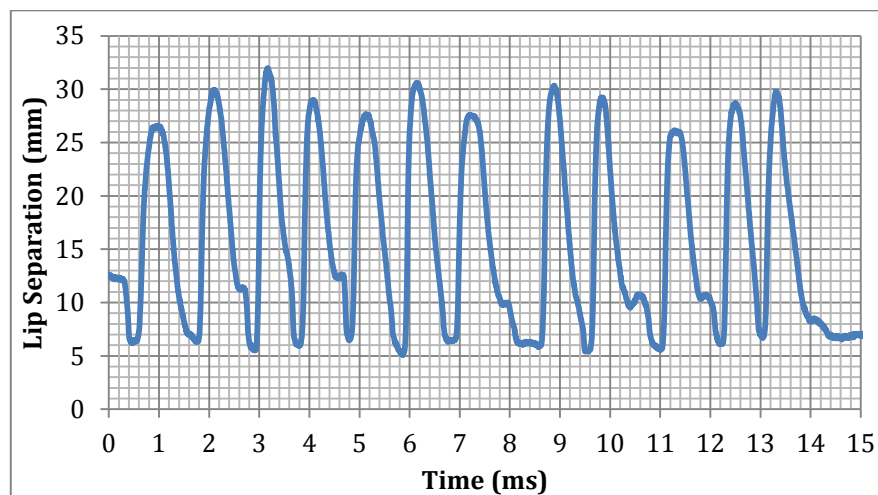


Figure 1. A typical plot of lip separation associated with 12 bilabial plosives produced by a male Cantonese speaker.

Results

Lip Opening Velocity (LOV)

Average LOV values (in mm/s) associated with all bilabial plosives produced at six Cantonese tones by the three speakers are shown in Figure 2a.

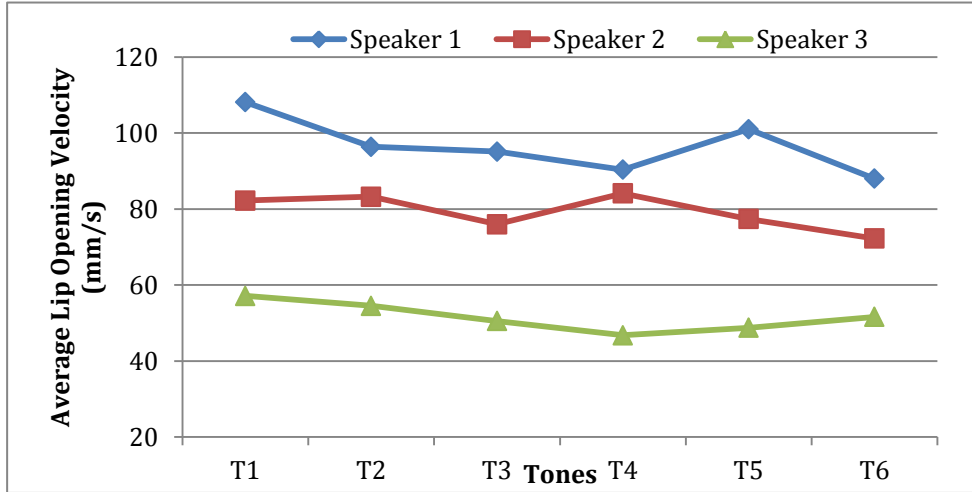


Figure 2a. Average LOV values (in mm/s) associated with all bilabial plosives produced at six Cantonese tones by the three speakers.

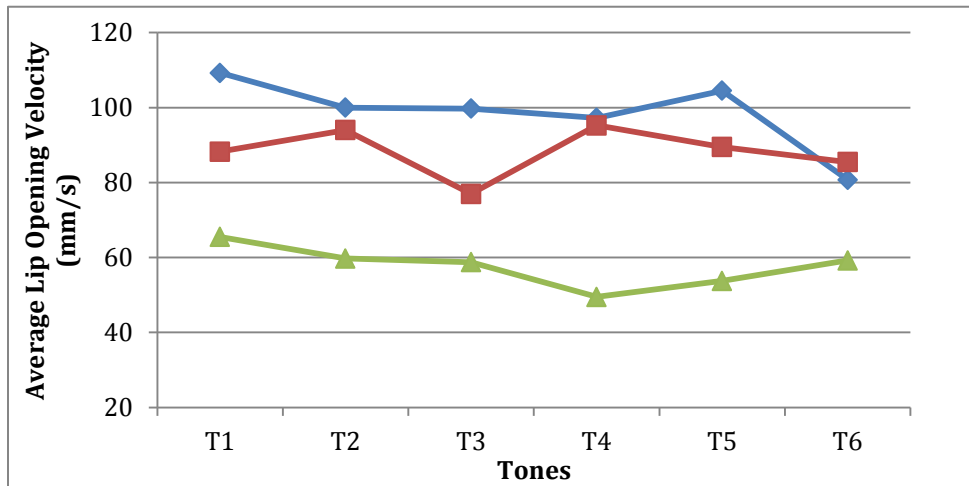


Figure 2b. Average LOV values (in mm/s) associated with aspirated bilabial plosives produced at six Cantonese tones by the three speakers.

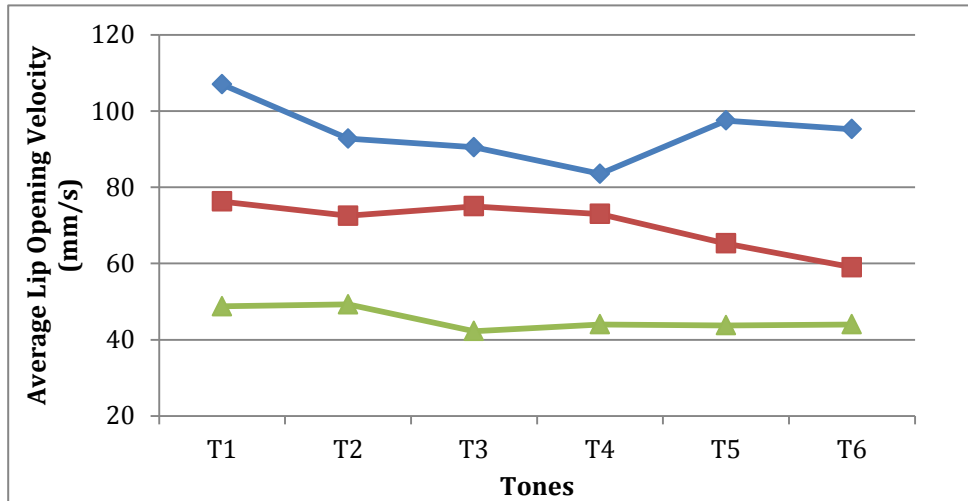


Figure 2c. Average LOV values (in mm/s) associated with unaspirated bilabial plosives produced at six Cantonese tones by the three speakers.

According to Figure 2a, a generally decreasing trend of LOV values was observed, with T1 having the greatest value and T6 having the smallest value. In fact, this was true for both aspirated and unaspirated plosives (see Figures 2b and 2c). For both aspirated and unaspirated productions, T5 exhibited the greatest fluctuations, with either a high or low LOV value, whereas other tones fluctuated only to a restricted range. In addition, such pattern of fluctuating LOV associated with T5 appeared to be more remarkable for unaspirated than that for aspirated productions.

Moreover, bilabial plosives were compared among the three speakers to reveal possible individual differences in LOV values. Average LOV values of the aspirated and unaspirated bilabial plosives produced at the six tones were collapsed and depicted in Figure 3. As shown in the Figure, the unaspirated plosives were associated with consistently higher average LOV values than the aspirated counterparts for all three speakers. Both Figures 2 and 3 collectively reveal the presence of considerable amount of inter-speaker variation. Generally speaking, Speaker 1 demonstrated the greatest LOV values and Speaker 3 had the smallest LOVs.

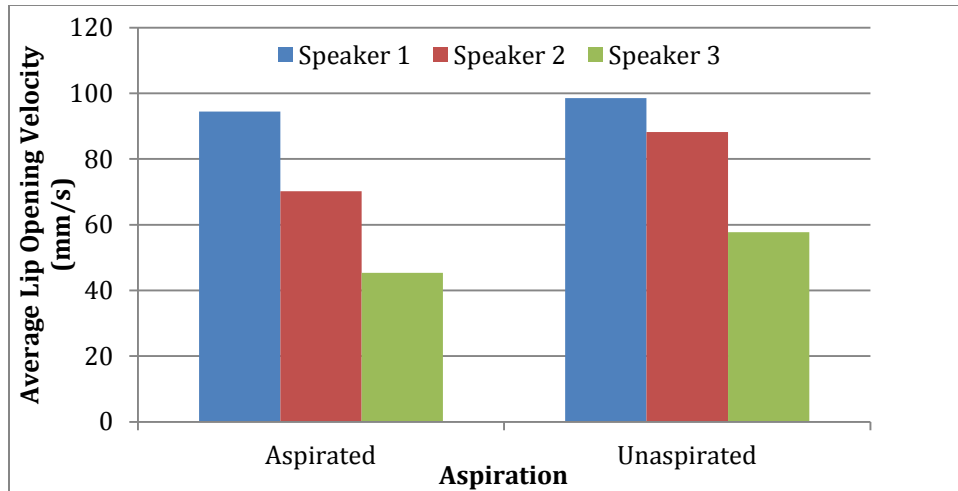


Figure 3. Average lip opening velocity (LOV) values associated with aspirated and unaspirated bilabial plosives produced by the three male Cantonese speakers.

Peak Velocity

Average peak velocity values associated with bilabial plosives produced at six Cantonese tones exhibited by the three speakers are shown in Figure 4a. Figures 4 b and 4 c depict the average peak velocity values associated with the aspirated and unaspirated bilabial plosives produced at the six Cantonese tones.

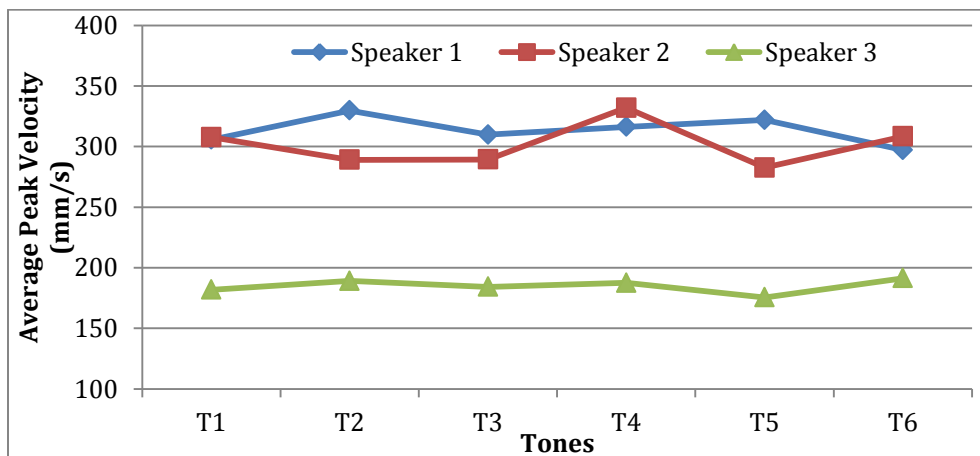


Figure 4a. Average peak velocity values of bilabial plosive produced at different tones by three male Cantonese speakers.

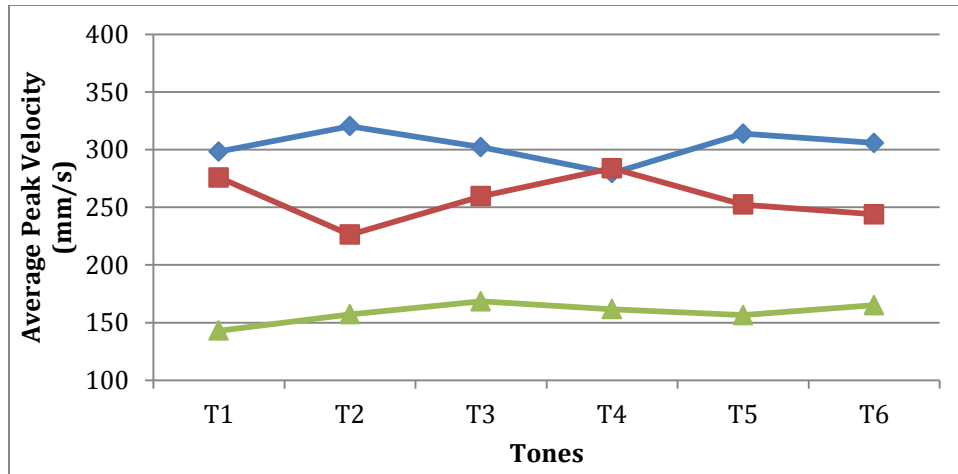


Figure 4b. Average peak velocity values of aspirated bilabial plosive produced at six Cantonese tones by the three male speakers.

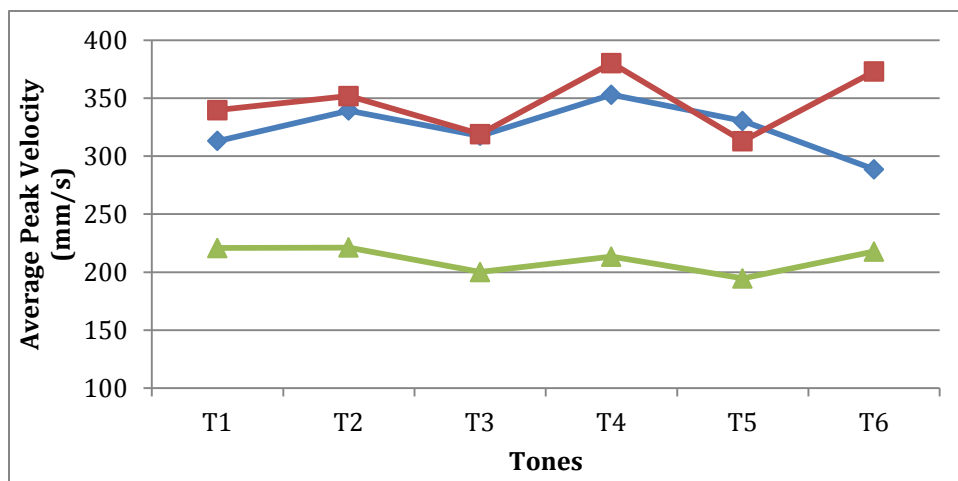


Figure 4c. Average peak velocity values of unaspirated bilabial plosive produced at six Cantonese tones by the three male speakers.

Plosives produced by the three speakers were compared in order to examine the possible effect of tones on lip movement. According to Figures 4b and 4c, a relatively constant peak velocity was noted for both aspirated and unaspirated plosives. No observable change in peak velocity value was found from T1 to T6. Though fluctuations in peak velocity value in different

tones were not consistent in aspirated and unaspirated productions, the variations in peak velocity value for T4 and T6 were notable. Also, there was a considerably larger variation in peak velocity value among tones for all speakers in unaspirated production.

To examine the individual differences, the peak velocity values of plosives produced by the three speakers were compared. Average peak velocity values of both aspirated and unaspirated bilabial plosives with six tone productions are shown in Figure 5. A consistently higher peak velocity value exhibited in unaspirated productions for all three speakers was shown in the figure. Furthermore, from both Figure 4 and 5, a noticeable inter-speaker variation was revealed as Speaker 1 and Speaker 2 having similarly large peak velocity values and Speaker 3 having the smallest peak velocity values.

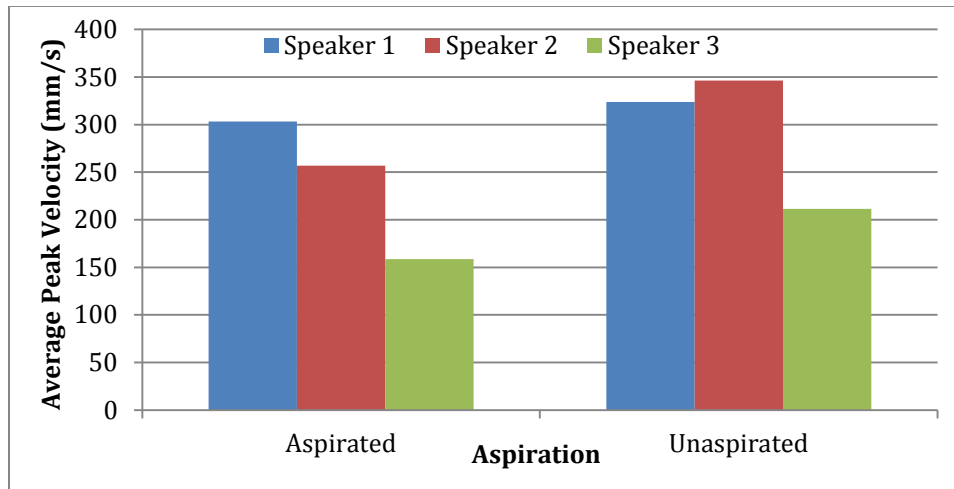


Figure 5. Average peak velocity values associated with aspirated and unaspirated bilabial plosives produced by the three male Cantonese speakers.

Time-to-peak velocity (TPV)

Average TPV values of bilabial plosives produced at six Cantonese tones by the three speakers are tabulated in Figure 6a. The average TPV values for aspirated and unaspirated plosives are shown in Figure 6b and 6c respectively.

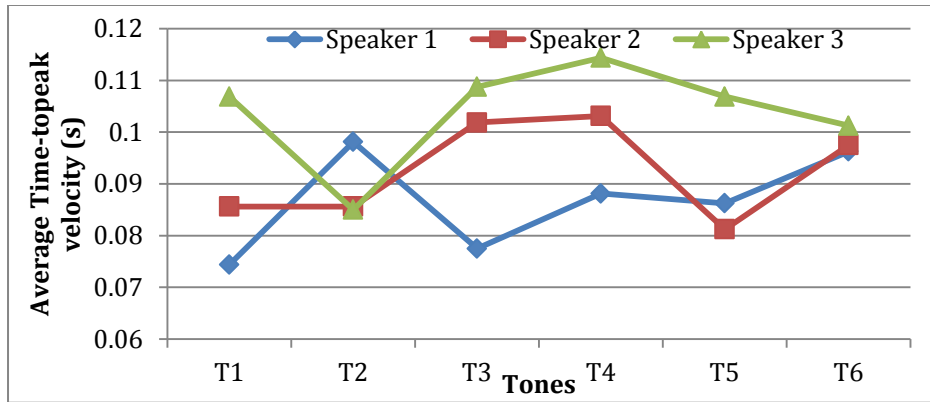


Figure 6a. Average TPV values of bilabial plosives produced at the six Cantonese tones by the three speakers.

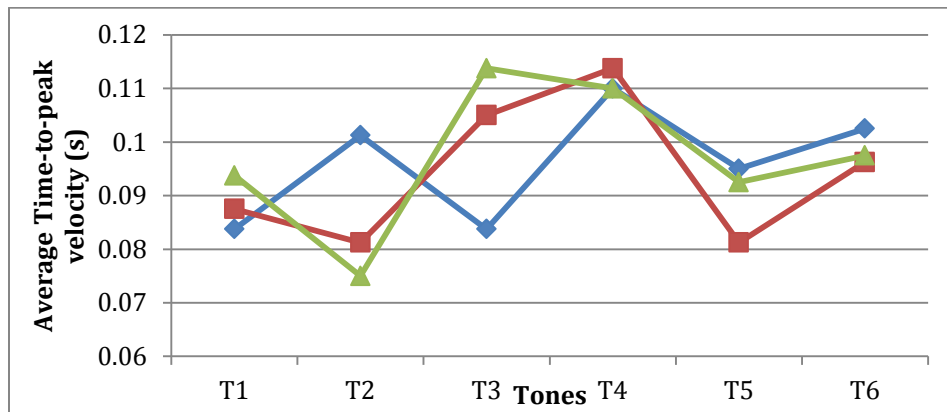


Figure 6b. Average TPV values of aspirated bilabial plosives produced at the six Cantonese tones by the three speakers.

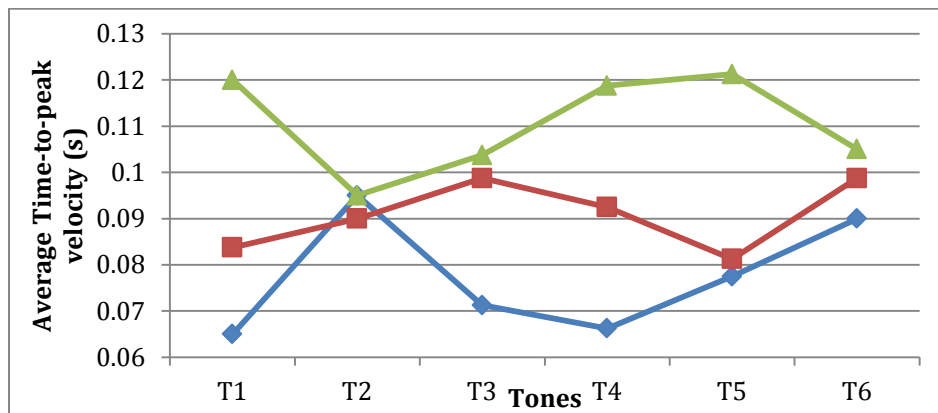


Figure 6c. Average TPV values of unaspirated bilabial plosives produced at the six Cantonese tones by the three speakers.

As shown in Figure 6a, a general trend of slightly increasing TPV across T1 to T6, with T1 having a smaller TPV than T6 was found. However, for both aspirated and unaspirated plosive, this pattern of TPV change was not obvious as variation in TPV was considerably large, with unaspirated plosives having a slightly higher degree of TPV variation. Observable differences in TPV values were found in T2, T3 and T5 which different speakers had produced the tones with either high or low TPV values.

In addition, individual differences in TPV values can be observed by comparing three speaker's productions. Average TPV values associated with aspirated and unaspirated plosives produced by the three speakers are displayed in Figure 7. In general, TPV values of aspirated plosives were slightly higher than those of unaspirated productions. However, a more consistent TPV values was found in aspirated plosives. From both Figures 6a and 7, a less fluctuation was observed in TPV values among speakers, especially for the aspirated production.

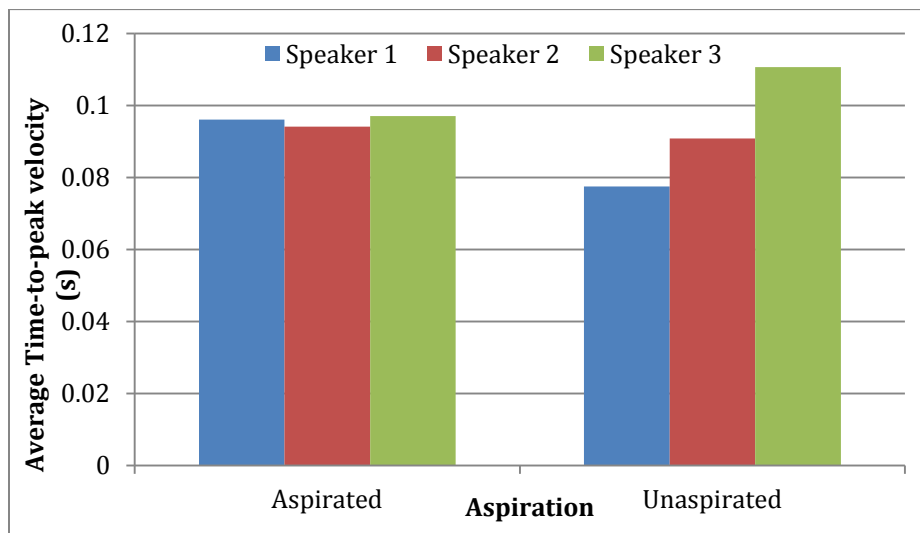


Figure 7. Average time-to-peak velocity (TPV) values associated with aspirated and unaspirated bilabial plosives produced by the three male Cantonese speakers.

Discussion

Tones

By comparing the average values of LOV, peak velocity of lip opening and TPV during the plosive production, the effect of tones on bilabial movement was investigated. LOV of T1 (high level tone) was consistently higher than that of T6 (low level tone), with a generally decreasing pattern across the tones (see Figure 2a). This implies that, for plosives produced at high tones, it took a shorter time to reach maximum lip separation than plosives produced at low tone. According to speech aerodynamics, high tones are produced a faster vocal fold vibration. In order to vibrate the vocal folds more rapidly, both longitudinal tension of the vocal folds and subglottal pressure are found to be higher (Titze, 1989; Plant & Younger, 2000). With higher pressure beneath the vocal folds, more air will be released during the burst phase, and the lips should open to the maximum separation at a faster rate as the lips are now under a stronger opening force. Therefore, plosives produced at a higher tone are associated with a faster lip opening movement than those produced at a lower tone.

According to Figure 4a, the peak velocity of lip opening movement associated with the plosives was generally constant across the six Cantonese tones. It follows that peak velocity of lip opening during bilabial plosive production appears to be independent of the lexical tone. As mentioned above, higher tone is associated with greater subglottal pressure. The constant peak velocity of lip opening found for all tones seems to suggest that, in addition to aerodynamic events, other factors such as other mechanical factors (e.g., the opening force of the mandible) may also play a role in determining the kinematics of lip opening. Apparently, further studies of the actual mechanics of lip and/or jaw opening during plosive production are necessary to precisely answer this question.

In addition, according to Figure 6a, TPV (time to peak velocity) values of bilabial plosive production of three speakers were found to generally increase from high tone (T1) to low tone (T6). It follows that it took a longer time for a low tone plosive to reach the maximum lip opening velocity than a high tone plosive. This is consistent with the findings of LOV across different tones. As the maximum velocity was achieved before the maximum opening, the greater was the TPV, the longer time it took to achieve maximum lip opening. Mathematically, average LOV could be obtained by dividing the maximum separation of lip during plosive production by the time elapsed, As TPV is inversely proportional with LOV, how tone effects TPV could be explained in the same way as LOV. As high tones are associated with higher subglottal pressure, a greater force is available to burst open the lips (Titze, 1989), implying a shorter amount of time required to reach the peak opening velocity (greater acceleration of lip opening). In summary, greater subglottal pressure that is associated with higher tone shortens the time required for the lips to reach maximum separation.

In conclusion, tone height effect on Cantonese bilabial plosive articulation was found. Articulatory movement of lips was generally faster in high tone plosives and low tone plosives. This finding was comparable with a kinematic study of Mandarin bilabial plosives, having difference in articulatory movement of jaw and tongue in high tone and low tone productions (Erickson et al., 2004). However, as the articulators studied were different, interpretation of results should be done with caution.

Inconsistency in tone production

According to the figures, some tones particularly T2, T4 and T5 were produced with high degree of variation or inconsistency. Such inconsistency in tone production may be due to speakers' difficulty in perception and production (Ciocca & Lui, 2003; Ip, 2006). Two types of

difficulty were addressed in low tones and rising tones. According to Ip's (2006) study of tone acquisition of Cantonese, difference in fundamental frequency (F0) among low tones was narrower than that of high tones. In other words, low tones are not as easily distinguishable as high tones, rendering a lower accuracy in producing low tones. A smaller F0 difference therefore requires more accurate motor control. Logically, it also yields greater chance of variation or inconsistency within and among speakers. As shown in the kinematic data in Figures 2, 4, and 6, greater amount of variations were seen in low tone productions which may have influenced data interpretation.

For perception and production of rising tones, Ciocca and Lui (2003), and Ip (2006) suggested that high-rising tone and low-rising tone was the poorest pair in discrimination task among young children. This was further supported by Wong, Ciocca, and Yung (2009) in examining the performance of the similar task by adults. In addition, due to the similarity of rising pattern in pitch contour and fundamental frequency onset, production of high-rising tone and low-rising tone was noted with a remarkably low accuracy compared to the accuracy of high-level tone and low-level production (Ip, 2006). Confusion between high-rising tone and low-rising tone was reported during the data collection of the present study. Moreover, such inaccurate and uncertain productions led to the variation within and among speakers, contributing to difficulty in interpreting the experimental results. With the considerable variations or inconsistencies, the present data may have been influenced. Improvements such as increasing the number of production tokens, number of participants and practice time is therefore recommended.

Aspiration

The effect of aspiration on bilabial movement was revealed by comparing the values of LOV, peak velocity of lip opening and TPV during production of aspirated and unaspirated plosives. As seen in Figure 3, average LOV values of unaspirated plosives were found to be consistently higher than those of aspirated plosives. Aspirated plosives required a longer time to achieve maximum lip separation. In comparing bilabial plosive productions in different languages, Butcher (1992) reported that aspirated plosive /p^h/ was consistently associated with a greater peak intraoral pressure than unaspirated plosive /p/. Intuitively, the greater peak intraoral pressure in aspirated plosives should be followed by faster lip movement during the burst phase. However, the present findings appear to contradict this suggestion; unaspirated plosives with lower intraoral pressure were found to have a faster lip opening, as indicated by the greater average LOV values. Yet, the finding of peak velocity of lip movement is in line with this; unaspirated plosives were consistently associated with higher peak velocity values than aspirated plosives. (see Figure 5a.) This suggests that the fastest lip movement during burst phase was recorded sooner in unaspirated plosives than that of aspirated plosives.

Lastly, the time-to-peak-velocity values of aspirated plosive and unaspirated plosive were compared and shown in Figure 7. The aspirated bilabial plosives took a longer time to reach the peak velocity than the unaspirated plosives, suggesting a greater acceleration in lip movement in aspirated plosives. This was observed in two out of the three speakers. It follows that aspirated plosives with higher intraoral pressure exhibited slower peak velocity and slower acceleration of lip opening. Unaspirated plosives had greater lip opening velocity, peak velocity and acceleration (derived from TPV) than aspirated plosives. This appears to contradict to the physical mechanism since higher intraoral pressure of aspirated plosives should provide a stronger force

for faster lip opening movement. In a similar kinematic study of whispered bilabial plosives /p/ and /b/ of English, Higashikawa et al. (2003) found that whispered /b/ was produced with significantly higher peak velocity in lip separation than whispered /p/, and no difference was noticed between voiceless /p/ and voiced /b/. The articulatory gesture of whispered /b/ and whispered /p/ are comparable to unaspirated /p/ and aspirated /p^h/ in Cantonese, respectively. Higashikawa et al. (2003)'s findings appeared to be in line with the present findings of greater peak velocity of lip opening in unaspirated plosives than aspirated counterparts. Higashikawa et al. (2003) also found that the peak airflow of whispered /p/ was higher than that of whispered /b/, which enhanced the production turbulence in aspiration. This might also be used to explain the current findings for Cantonese aspirated plosives. In addition to such aerodynamic events, a hypothesis about motor coordination was evoked by the difference in kinematics of lip movement of two whispered plosives. Higashikawa et al. (2003) proposed that there were some difference in the motor coordination between whispered /b/ and /p/ in order to provide a better discrimination of two productions even without voicing properties. According to Higashikawa et al. (2003), increased lip aperture in whispered /b/ might be a way to reduce the power of high frequency noise and turbulence that found in aspirated feature in bilabial /p/ production. A similar notion was previously suggested by Gracco (1994) regarding the difference between voiced and voiceless sound pairs in terms of time phase for better accommodation of laryngeal voicing or devoicing. According to Gracco (1994), voiceless /p/ was recorded to have a faster lip closing velocity for lip and laryngeal activities while slower lip opening velocity for formulating an appropriate voice onset time. This further supports the finding in present study that greater velocity and acceleration of lip opening were associated with unaspirated Cantonese /p/, which is similar in articulatory gesture and aerodynamic conditions to the voiced /b/ in English.

Based on the findings from the present study and supportive evidence from the study of Higashikawa et al. (2003), two possible explanations are suggested. First, as only the vertical z-movement was studied in the present study, it may be risky to ignore the effect of high intraoral pressure on the peak velocity of lip movement. Seeing the widening of lip opening observed in Higashikawa et al.'s study, it is possible that the increased intraoral pressure only contribute a little to the peak velocity in the z-dimension. It may contribute more significantly to the x-dimension (anteroposterior dimension) of lip movement. A comprehensive kinematic study should be carried out to compare kinematic changes in three axes movement of aspirated and unaspirated plosive production. Secondly, a hypothesis originated from the findings of high intraoral pressure and low peak velocity value in aspirated plosive production, which was reported in Gracco (1994) which recorded the fastest lip opening in bilabial /m/ comparing with /b/ and /p/ despite /m/ was believed to have the lowest intraoral pressure. As there are voiced plosives in English, results of the study may not be directly generalizable to Cantonese. Yet, it provides some valuable insight for revealing other factors that may influence the articulatory kinematics during production of Cantonese plosives. It is possible that intraoral pressure may not be the only determinant for the specific aerodynamics and lip movement of Cantonese bilabial plosives. Two driving forces may be responsible for lip opening in aspirated and unaspirated plosives: aerodynamic force and muscular force. Lip opening can be facilitated by either aerodynamic force (lip closure being blown open by the great intraoral pressure) or the muscular force (mandibular depressor such as lateral pterygoid muscle lowering the jaw and thus separating the lips). It is speculated that aerodynamic force is mainly responsible for the lip opening in aspirated plosives, and muscular force is for unaspirated plosives. This is apparently a

conjecture. Further study both aerodynamics and physiology of lip opening during plosive production is needed to confirm this suggestion.

Individual variations

As discussed in the literature of lip kinematics, individual variability is common and this is believed to have significant influence on data analysis (Lofqvist & Gracco, 1997; Higashikawa et al., 2003). Consistently, significant variation is found in the present study, as indicated by differences in values, range of values and pattern of values were found obvious among three speakers. The pattern of inter-speaker variation found in the present study appears to be unpredictable. In fact, some kinematic values varied in a reversed manner (see Figures 5b and 6c). Alongside with small the number of speakers, interpretation of data has become difficult and the representativeness of the results is limited.

Limitations

Non-word familiarity

With the three non-word stimuli (two aspirated and one unaspirated plosives), the participants needed to practice for five minutes in order to familiarize themselves with novel syllables. Yet, accuracy of these non-words was not as high as real words as articulation of non-words probably was not natural. This may have limited the accuracy of the kinematic measurements in the present study.

To reduce the artifact associated with the use of non-words, it is suggested that longer practice time be given to the participants until a natural production is achieved. This can increase the naturalness of the lip movement and hence a more reliable kinematic measurement can be obtained.

Individual variation

Individual variation greatly limits the representativeness of the results. This can be improved by recruiting more participants. Moreover, as intra-speaker variation was also noticed in four productions of each word, this also reduced the accuracy of kinematic data. To reduce the effect of intra-speaker variation, increasing the number of productions is a possible solution. Wong, Schwartz, and Jenkins (2005) suggested that producing more than 240 tokens of the same syllable by the same speaker is likely to yield a more accurate and reliable results.

Further investigation

As discussed above, with a few modifications and additional measurements, the present study can be improved and a more reliable and representative results can be obtained.

Participants and materials

As discussed above, the huge individual variation can be controlled by having more participants producing more speech tokens. Moreover, longer practice period can be provided to the participants. This can reduce the unfamiliarity of non-words and the confusion between T2 and T5 productions.

Aerodynamic study

Since it is hypothesized that the kinematic control of lip movement is under the influence of both bilabial muscular activities and aerodynamic properties, the amount of airflow and subglottic pressure will be the important data revealing the laryngeal behaviors during speech production. Aerodynamic measurement can facilitate the result analysis of tone effect and aspiration effect on bilabial plosives.

Lip closing kinematics

As claimed by Gracco (1994), lip closing velocity was recognized as an important parameter for disclosing the effects of bilabial muscular activities since lip closure was more voluntarily done and had less affected by the oral pressure. Lip closing kinematics can also provide more information about the preparation of plosive production (Gracco, 1994; Higashikawa et al., 2003).

As found that there were differences in kinematic data for lip opening in tones and aspirated/unaspirated productions, study of lip closure can provide further experimental data in investigating the effect of tones and aspiration.

Conclusion

In present study, the effect of aspiration and tone on the lip movement during bilabial plosive /p^h/ and /p/ in Cantonese was examined by using electromagnetic articulography (EMA). By measuring lip opening velocity (LOV), peak velocity of lip opening and time to peak velocity (TPV) from three male participants, comparison was made for the kinematics of lip movement during production of aspirated and unaspirated plosives at the six Cantonese.

Generally speaking, lip opening was found to be faster in plosives produced at high tones than those produced at low tones, as indicated by the generally decreasing trend of velocity and acceleration noticed from T1 to T6. This was coherent with the tone height study of tongue and jaw movement in Mandarin (Erickson et al., 2004). A possible explanation was related to higher subglottic pressure required for high tones, yielding a faster lip movement. For the effect of aspiration, unaspirated plosives were associated with higher peak velocity and acceleration during lip opening than the aspirated counterparts. Further studies especially involving

aerodynamic measurement and laryngeal activities should be done before a more concrete conclusion can be drawn.

In addition, individual variability and inaccuracy in low and rising tone productions had contributed to a potential limitation of data interpretation. Probably due to the high inter-speaker variation in speaking style, the present kinematic data was not consistent in the values and patterns. Moreover, as low tones (T4 - T6) had a narrower range in fundamental frequency than high and mid tone (T1- T3), greater variation of lip movement during production was recorded. Similarly, for rising tone (T2 and T5) productions, lip movement was also found to be less consistent. With these inconsistencies, interpretation of data needs to be carried out with caution and representativeness data seems limited. Further investigations can be done with focuses on different aspects including the aerodynamics, three-dimensional study of lip movement, lip closing movement and contextual variation.

References

- Bartle, C. J., Goozée, J., Scott, D., Murdoch, B. E., & Kuruvilla, M. (2006). EMA assessment of tongue-jaw co-ordination during speech in dysarthria following traumatic brain injury. *Brain Injury, 20*(5), 529-545
- Butcher, A. R. (1992). Intraoral pressure as an independent parameter in oral stop contrasts. In J. Pittam, *Proceedings of the 4th Australian International Conference on Speech Science and Technology*, 286–291
- Ciocca, V. & Lui, J. Y. K. (2003). The development of the perception of Cantonese lexical tones. *Journal of multilingual communication disorders, 1*(2), 141-147
- Clark, J., Yallop, C., & Fletcher, J. (2007). *An Introduction to Phonetics and Phonology*. Oxford: Blackwell Publishing
- Erickson, D., Iwata, R., Endo, M., & Fujino, A. (2004). Effect of tone height on jaw and tongue articulation in Mandarin Chinese. *Proceedings of the International Conference on Tonal Aspects of Languages*, Beijing, China
- Fant, G. (1960). *Acoustic Theory of Speech Production*. The Hague: Mouton
- Goozée, J., Murdoch, B. E., Ozanne, A., Cheng, Y., Hill, A., and Gibbon, F. (2007). Lingual kinematics and coordination in speech-disordered children exhibiting differentiated versus undifferentiated lingual gestures. *International Journal of Language and Communication Disorders, 42*(6), 703-724
- Goozée, J., Murdoch, B.E., Theodoros, D.G., & Stokes, P.D. (2000). Kinematic analysis of tongue movements in dysarthria following traumatic brain injury using electromagnetic articulography. *Brain Injury, 14*(2), 153-174

- Gracco, V. L. (1994). Some organizational characteristics of speech movement control. *Journal of Speech and Hearing Research, 37*, 4-27
- Higashikawa, M., Green, J. R., Moore, C. A., & Minifie, F. D. (2003). Lip kinematics for /p/ and /b/ production during whispered and voiced speech. *Folia Phoniatica et Logopaedica, 55*, 17-27
- Hoole, P., & Hu, F. (2004). Tone-vowel interaction in Standard Chinese. *Proceedings of the International Conference on Tonal Aspects of Languages, Beijing, China*
- Hoole, P., Zierdt, A., & Geng, C. (2003). Beyond 2D in articulatory data acquisition and analysis. *Proceedings of the International Conference of Phonetic Sciences, XVth*, pp. 265-268
- Ip, W. V. (2006). *Perception and production of lexical tones by young children*. Dissertation
- Löfqvist, A. (1996). Control of oral closure and release in bilabial stop consonants. In P. McCormack & A. Russel (Eds.), *Proceedings of the Sixth Australian International Conference on Speech Science and Technology*, pp. 561-566. Canberra: The Australian Speech Science and Technology Association.
- Lofqvist, A., & Gracco, V. L. (1997). Lip and jaw kinematics in bilabial consonant production. *Journal of Speech, Language, and Hearing Research, 40*, 877-893
- Ng, M., & Chu, R. (2009). An acoustical and perceptual study of vowels produced by alaryngeal speakers of Cantonese. *Folia Phoniatica et Logopaedica, 61*, 97-104
- Perkell, J. S., Cohen, M. H., Svirsky M. A., Matthies, M. L., Garabeta, I., & Jackson, M. T. T. (1992). Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements, *Journal of the Acoustical Society of America, 92*(6), 3078-3096
- Pickett, J. M. (2001). *The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory, and Technology*. Boston, MA: Allyn and Bacon.

- Plant, R. L., & Younger, R. M. (2000). The interrelationship of subglottic air pressure, fundamental frequency, and vocal intensity during speech. *Journal of Voice, 14*(2), 170-177
- Titze, I. R. (1989). On the relation between subglottal pressure and fundamental frequency in phonation. *Journal of the Acoustical Society of America, 85*(2), 901-906
- Wong, P., Schwartz, R. G. & Jenkins, J. J. (2005). Perception and Production of Lexical Tones by 3-Year-Old Mandarin-Speaking Children. *Journal of Speech, Language, and Hearing Research, 48*, 1065–1079
- Zierdt, A., Hoole, P., & Tillmann, H. G. (1999). Development of a system for three-dimensional fleshpoint measurements of speech movements. *Proceedings of the XIVth International Congress of Phonetic Sciences, 73-76*

Acknowledgment

I would like to express wholehearted gratitude to Dr. Manwa L.Ng from Division of Speech and Hearing Sciences of The University of Hong Kong for his advice and close supervision. I would also like to thank Mr. Dustin Wang for his technical assistance for the use of the EMA during data collection. Also, I wish extend my sincere thanks to my classmates and friends, particularly Andy, Eric, Florence, May, Morris and Timothy for their support and help. Finally, I would like to give a special thanks to my family for their everlasting support.

Appendix A

Consent Form

Interaction between lexical tone and labial movement in Cantonese bilabial

plosive production

You are invited to participate in a research study conducted by Mr. Eric Tong Tik Sang, a year four student, supervised by Dr. Manwa Lawrence Ng in the Division of Speech and Hearing Sciences at the University of Hong Kong.

PURPOSE OF THE STUDY

The study aims to investigate the relationship of lip movements and the tonal change in Cantonese by 3-dimensional Electromagnetic Articulograph (EMA) (AG500, Carstens Medizinelektronik)

PROCEDURES

You will be invited to participate in this study by reading out a list of Chinese words in Cantonese in an EMA cube. Your movements of lips and other oral parts will be visually tracked by EMA and your speech will be recorded by a microphone. Before and during the study, six tiny electromagnetic coils attached on your lip surface (upper lip and lower lip), the jaw, the bridge of nose, the back of an ear and the gums above two upper incisors using biologically safe adhesive. All the procedures will be done in the Speech Science Laboratory, 5/F Prince Philip Dental Hospital; it will take you about 20 to 30 minutes.

POTENTIAL RISKS / DISCOMFORTS AND THEIR MINIMIZATION

The EMA device will emit an endurable strength of electromagnetic field from a cube which is biologically harmless to human. Please inform the investigator if you have hypersensitivity to electromagnetism, fitting of pacemaker and claustrophobia, for any doubts about the possible reactions to the electromagnetism, please feel free to inquire the investigator.

In addition, the affixation of electromagnetic coils on your lips, gums and nose may cause a certain degree of discomforts which are no greater than what we experience in everyday life.

COMPENSATION FOR PARTICIPATION

You will receive no compensation for your participation. Your help will be deeply appreciated.

POTENTIAL BENEFITS

There are no direct benefits to you. However, the study can provide valuable information on the relationship of lip movements and tonal changes in Cantonese. This information in turn could provide additional information in review of the source-and-filter theory.

CONFIDENTIALITY

All of your information (kinematic and acoustic data of your speech and articulatory movements) obtained in the study will be kept under strict confidentiality. Data will be only assessable to the investigators for research purposes only. Complete destruction of the data will be done after the completion of the study.

PARTICIPATION AND WITHDRAWAL

Your participation is voluntary. This means that you can choose to stop at any time without negative consequences.

QUESTIONS AND CONCERNS

If you have any questions or concerns about the research, please feel free to contact Mr. Eric Tong Tik Sang at 93481550 and tongtiksang@gmail.com and the research supervisor, Dr. Manwa Lawrence Ng at 28590582 and manwa@hku.hk. If you have questions about your rights as a research participant, contact the Human Research Ethics Committee for Non-Clinical Faculties, HKU (2241-5267).

SIGNATURE

I _____ (Name of Participant)

understand the procedures described above and agree to participate in this study.

Signature of Participant Date

Appendix B

運動學研究: 廣東話雙唇音聲調與嘴唇運動之相互影響

香港大學言語及聽覺科學部現正邀請您參與由四年級學生湯迪生主理並由吳民華博士指導的研究調查。

研究目的

是次研究希望利用三維電磁構音儀 (Electromagnetic Articulograph, EMA) (AG500, CarstensMedizinelektronik) 就廣東話雙唇音聲調與嘴唇運動的關係作出調查。

研究程序

您將需要坐在電磁構音儀內，然後以廣東話讀出所附列表上的文字。您的嘴唇及其他面部及口肌運動將被電磁構音儀紀錄，另外，您的朗讀過程亦會被錄音。在研究開始及進行時，六個細小的電磁圈將會以對人體無害的膠粘劑分別粘貼在您的上下唇、顎、鼻樑及上門牙的牙肉上。

整個研究過程將會在香港大學菲臘牙科醫院言語及聽覺科學部的言語科學實驗室內進行，為時二十至三十分鐘。

潛在風險

電磁構音儀會發放出一定強度的電磁波，經測試絕對為人體可承受並無任何副作用產生。若您對電磁波敏感、安裝了心臟起搏器或有幽閉恐懼症而未能坐在電磁構音儀內進行研究，請通知研究員。您亦可就著有關對電磁波可能所產生的反應向研究員提出疑問。

另外，電磁圈在粘貼在您的上下唇、顎、鼻樑及上門牙的牙肉上之後會有一定程度但絕不會多於日常生活所體驗的不適。

參與回饋

感謝您的參與，是次研究活動並不會提供任何物資或金錢上的回饋。

可能利益

是次研究不會為您帶來利益。但你將會為是次研究課題提供了有用的資料，並為將來就音源過濾理論(source-and-theory)的審視作出幫助。

私隱

包括您的語音運動表現及錄音記錄在內，在研究中所獲取的資料將會絕對保密。所有資料只有研究員以研究目的使用，並在研究完畢後銷毀。

參與及終止

由於您的參與屬自願性質，故此你可以於研究的任何時間選擇終止而不會引致任何不良後果。

如您對是項研究有任何問題，可於現在提出。

如日後你對是項研究有任何查詢，請與湯迪生（電話號碼：93481550；電郵：tongtiksang@gmail.com）或是次研究導師吳民華博士（電話號碼：28590582；電郵：manwa@hku.hk）聯絡。如你想知道更多有關研究參與者的權益，請聯絡香港大學非臨床研究操守委員會（電話號碼：22415267）。

本人 _____ (參與者姓名)明白以上內容，並願意參與是項研究。

參與者簽署

日期

Appendix C

Word List of Cantonese Bilabial Plosives

巴

罷

霸

牛 (扒)

/p^ha5/

怕

爬

把

(爸) 爸

/p^ha6/

/pa5/

趴

霸

/p^ha5/

牛 (扒)

巴

趴

怕

/pa5/

爬

把

/p^ha6/

罷

(爸) 爸

/p^ha6/

(爸) 爸

趴

怕

把

爬

/p^ha5/

巴

罷

/pa5/

霸

牛 (扒)

爬

巴

罷

把

(爸) 爸

/p^ha6/

/pa5/

怕

牛 (扒)

霸

趴

/p^ha5/

/p^ha5/

怕

罷

把

爬

趴

/p^ha6/

/pa5/

巴

(爸) 爸

牛 (扒)

霸